ROBUST STILL IMAGE CODING USING LAPPED TRANSFORMS WITH BLOCK CLASSIFICATION

Arthur L. A. da Cunha[†], *Weiler A. Finamore*[†], *and Eduardo A. B. da Silva*[††]

[†]Center for Telecommunication Studies
Pontifical Catholic University of Rio
Rua M. S. Vicente 255, Rio de Janeiro, Brazil
{arthur,weiler}@cetuc.puc-rio.br

[††]Universidade Federal do Rio de Janeiro
PEE/COPPE/DEL/EE
Rio de Janeiro, RJ - Brazil
eduardo@lps.ufrj.br

ABSTRACT

Very efficient still image compression methods exist today, but most of them are not fit for transmission on error prone channels. One example of state-of-the-art robust image coding is the use of wavelet compression followed by forward error correction, as proposed by Sherwood and Zeger, that gives high coding efficiency but suffers from the incomplete decoding problem. An alternative to this is the joint source/channel coding scheme by Cheng and Fischer, which overcomes this difficulty at the expense of a lower PSNR performance. In this paper we propose a joint source/channel coding scheme with increased coding efficiency based on lapped transforms with block classification. With an efficient bit allocation strategy and the high coding gain of Lapped Transforms, a significant increase, on the order of 2 dB, in the PSNR versus BER performance has been obtained over the classical joint source/channel scheme on binary symmetric channels.

1. INTRODUCTION

Digital imagery deals with large amounts of information that often need to be compressed and coded in order to be reliably transmitted over bandlimited channels. Shannon separation principle from information theory guarantees error free transmission by a source encoder in *tandem* with a channel encoder. However this holds only for arbitrarily long delay and complexity — two commodities not abundant in practical situations. A different approach, with practical appeal, is joint source/channel coding. Several joint source/channel coding schemes have been proposed for still images [1, 2, 3, 4]. Most of them are either a combination of channel optimized quantizers and efficient error-correcting codes, or channel-optimized quantizers alone. The latter has the advantage of being simpler since it dispenses with the use of channel codes.
Channel-optimized vector quantizer [5] and its scalar version, the channel-optimized scalar quantizer (COSQ) [6], are efficient approaches to source/channel coding. For memoryless Gaussian sources with mean square error distortion criterion, the performance of channel optimized quantizers, for moderate block length, is shown [5] to be better than that of their tandem counterparts. With other sources however, their performances are substantially away from $D(C)$, the distortion-capacity bound. In addition, the quantizer is sensitive to impulsive noise caused by wrongly delivered quantizer indices, due to channel errors. While these errors do not strongly affect the mean square distortion measure, they produce annoying salt-and-pepper noise-like artifacts.

In [3], a technique to cope with the impulsive noise problem, uses an efficient all-pass filtering with a binary phase scrambling/descrambling method followed by COSQ (we will refer to this method as phase scrambling). There the authors show that the performance for a wide class of Generalized Gaussian sources is equivalent to that of memoryless Gaussian sources quantized with COSQ. Moreover, the phase scrambling copes well with the salt-and-pepper artifact caused by the impulsive noise by spreading it, after decoding, all over the given subband.
In [2] a similar approach is obtained with the replacement of the COSQ by the channel-optimized TCQ (CO-TCQ) [7]. The increased performance thus obtained is a result of the long-delay decision search inherent to TCQ.

This paper is organized as follows. In Section 2 we describe the proposed scheme. In Section 2.1 we briefly review LT's and describe the bit allocation strategy. A subband/block classification scheme which significantly improves the overall performance is explained in Section 3. Results and conclusions are presented in Sections 4 and 5.

2. THE PROPOSED SCHEME

The block diagram of Fig. 1 displays the proposed scheme. The image samples which are first transformed by the selected Block Transform are directed to the Block Classification where the coefficients variances are estimated. It is assumed that the coefficients are, except for the DC, zero mean and approximately Laplacian distributed. The coefficient variances, bit allocation matrix, classification map, as well as the image mean are transmitted as side information. Since the amount of side information represented is small, with low overhead a good FEC scheme can be used to encode it, just as suggested in [3]. The bit allocation is accomplished by using the computed variances and the method indicated in [8].
The decoder simply follows the reverse process: the samples recovered from the received bits are de-normalized and the values obtained from the quantizer indices enter the phase descrambler, the output of which inputs the transform synthesis stage.

This work has the support of to CNPq (Conselho Nacional de Pesquisa e Desenvolvimento).

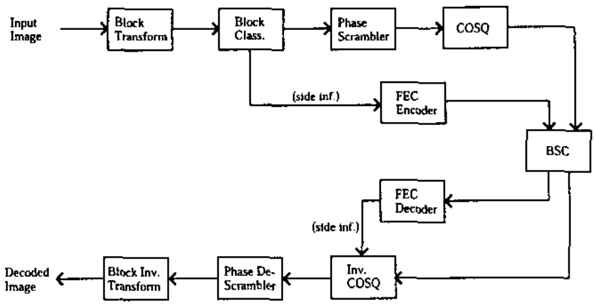

Fig. 1. Proposed Scheme Building Blocks

2.1. Lapped Transforms

Lapped transforms were first developed to reduce blocking artifacts inherent to block transforms such as the DCT. Research effort in recent years has led to efficient LT's optimized for high coding gain and thus superior PSNR performance. The existing fast algorithms allow small complexity implementations. In this work we investigate, for comparison reasons, compression schemes with the following LT's: (LOT) Malvar's Lapped Orthogonal Transform [9]; (G-LOT) de Queiroz' Generalized Orthogonal Lapped Transform [10]; (G-BLT) Tran et al's [11], 8 × 16, Generalized Lapped Bi-orthogonal Transform; (FLT) Fast multiplierless LT's [12].

As can be seen in Fig 2, the G-BLT is the one that provides the best PSNR versus BER results.

2.2. Robust Quantization

The coding performance achieved by the COSQ can be further improved by means of phase scrambling the subbands [2, 3]. The phase scrambling is accomplished by adding a reference random sequence to the phase of the input sequence. That is, after the Fast Fourier Transform (FFT) of each sequence is computed, the phase of the reference sequence is added to the phase of the input sequence followed by an inverse FFT.

The purpose of the phase scrambling is two-fold: (1) to cope with the impulsive noise induced by the channel, and (2) to obtain nearly Gaussian samples at the output of the filter in order to improve the quantizer performance. The scrambling can be used in conjunction with any quantization scheme. Our choice of COSQ stems from its low-computational burden in comparison to more sophisticated and computational-intensive schemes such as CO-TCQ.

3. BLOCK CLASSIFICATION AND BIT ALLOCATION

In order to improve the performance, especially at low bit rate, we have investigated a simple classification scheme. Since images are often characterized as non-stationary processes, significant gains can be obtained by allocating more bits to regions with more activity and therefore of greater importance.

The algorithm [13] is as follows. For each block i out of L image blocks, we first compute its AC energy

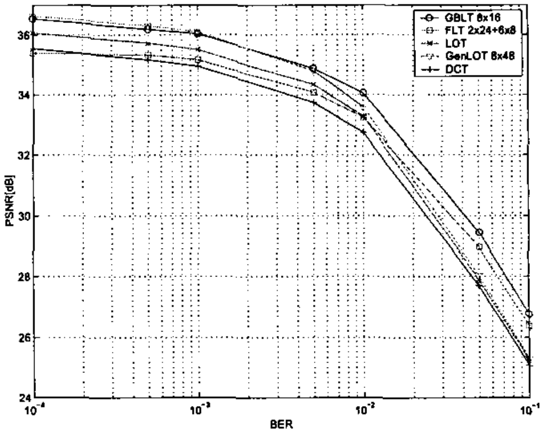

Fig. 2. Performance Comparison Between Various LT's and the DCT.

$$E_i = \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} G_i^2(k,l) - G_i^2(0,0) \tag{1}$$

and the block gain $g_i = \sqrt{E_i}$.

After the blocks are arranged in increasing order of block gain, we partition the blocks into J classes. Partitioning is done with respect to a set of thresholds $\{T_0 = 0, T_1, \ldots, T_J, T_{J+1} = \infty\}$ each block assigned to class j if the block gain lies between thresholds T_{j-1} and T_j.

The threshold selection seeks statistical homogeneity within each class. We use the algorithm presented in [13] to search for thresholds such that the ratio σ_i/m_i is kept constant, where σ_i and m_i are respectively the standard deviation and average for class i.

After block classification, bit allocation takes place. The problem of bit allocation can be understood as the selection of a set $\mathcal{R} = \{r_1, r_2, \ldots, r_N\}$ of bit rates that minimizes the overall distortion

$$D(\mathcal{R}) = \frac{1}{N} \sum_{i=1}^{N} d_i(r_i) \tag{2}$$

constrained to $\sum_{i=1}^{N} r_i \leq \bar{r}N$. In (2) $d_i(r_i)$ is the average distortion incurred by the i-th quantizer and $\bar{r}$ denotes the target average bit rate. An algorithm similar to one described in [8], a steepest descent method, that achieves an optimal allocation under certain conditions was used in the proposed scheme. This algorithm which is specially suitable for channel-optimized quantizers, can be summarized as follows:

1. Set $k = \bar{r}N$; Set $r_i = 0, i = 1, \ldots, N$
2. Set $k = k - 1$; find i_k which satisfies

$$\Delta_{i_k}(r_{i_k}) = \max_{i=\{1,\ldots,N\}} d_i(r_i) - d_i(r_i + 1) \tag{3}$$

3. Set $r_{i_k} = r_{i_k} + 1$. if $k = 0$ stop; else go to step 2.

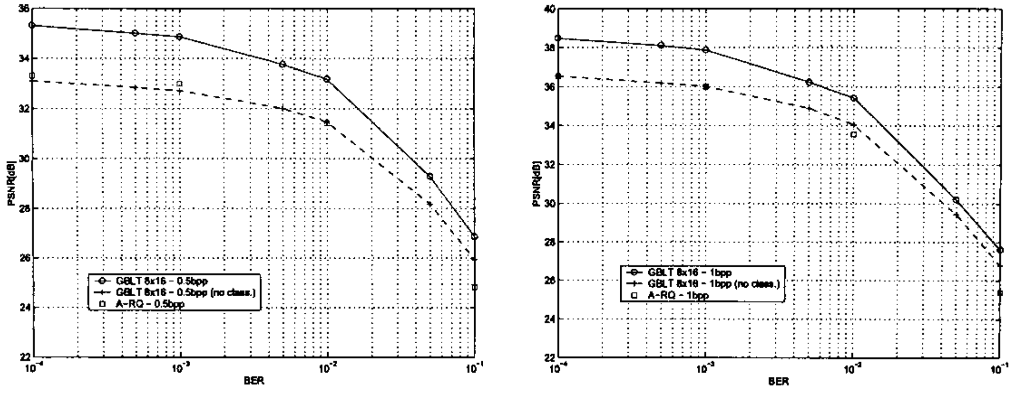

Fig. 3. Proposed scheme performance for Lena image at (a) 0.5bpp and (b) 1bpp.

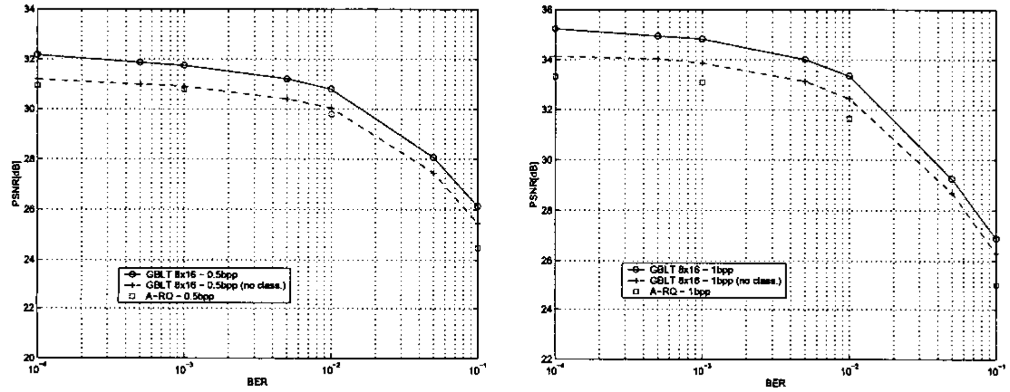

Fig. 4. Proposed scheme performance for Goldhill image at (a) 0.5bpp and (b) 1bpp.

We verified in our experiments that the algorithm produces results very close to an optimal allocation. Furthermore, with precomputed distortion-rate pairs the allocation is accomplished with a minimal number of operations and indeed runs very fast.

4. SIMULATION RESULTS

The performance of the proposed scheme was examined empirically by using Lena and Goldhill 512 × 512 pixels images. All the results shown are an average over ten experiments. The PSNR versus BER curves shown in Fig. 2 illustrates the performance of the scheme (*without classification*) at 1bpp for several choices of LT's. Also in the figure, for the sake of comparison, is the A-RQ wavelet scheme of Chen and Fischer. While the observed improvement for Lena image is small, for Goldhill image, improvements (more noticeable at high cross-channel probabilities) of up to $1.3dB$ were attained by our simulations. We believe that these improvements especially at low BER are due to the better bit allocation strategy of the proposed scheme and also, to the excellent synthesis performance of the bi-orthogonal LT.

In Fig's 3 and 4 the PSNR versus BER curves of the schemes (*with classification*) for 0.5 bit per pixel and 1.0bpp are displayed. Also shown in the figures are the scheme without classification (dotted lines). For a smooth image such as Lena, the curves show an improvement of up to 2.2dB at low BER. The same improvement was not noted as for the Goldhill image, but yet there is a 1dB improvement at moderate BER. In addition, subjective evaluation of the reconstructed images have shown better decoded images. As can be seen in Fig. 5, the Lena image at 0.25bpp generated using the scheme with classification is substantially better.

5. CONCLUSIONS

We have proposed a joint source/channel coding algorithm using robust quantization, Lapped Transforms and block classification. To our knowledge, the best published results in PSNR performance

(a) (b) (c)

Fig. 5. Example of images encoded at 0.25bpp over BSC with $P_e = 10^{-3}$ (a) Original Lena Image. (b) Prop. scheme without class. PSNR = 31.58dB (c) Prop. scheme with class. PSNR = 29.44dB.

are those of Sherwood and Zeger in [14]. Such schemes, however, usually are designed for a time-invariant channel and in addition they have a small probability of incomplete decoding. Our results narrow the gap between the performance of schemes such as Chen and Fischer which are not victimized by the incomplete decoding problem and that achieved by Sherwood and Zeger. We observed an improvement of over 2 dB of the proposed scheme with classification over Chen and Fischer's. We also note that the classification brings a significant contribution to the improvement on the proposed method performance. Our scheme has also the advantage of being robust to channel mismatch and does not suffer the incomplete decoding problem. In addition it does not explicit use channel codes, lowering the overall complexity.

6. REFERENCES

[1] S.S Channappayya, G. P. Abousleman, and L. J. Karam, "Joint source-channel coding of images using punctured convolutional codes and trellis-coded quantization," *Int. Symp. Circ. and Sys.(ISCAS)*, 2001.

[2] Tuyet-Trang Lam, Glen P. Abousleman, and Lina J. Karam, "Image coding with robust channel-optimized trellis-coded quantization," *IEEE JSAC*, vol. 18, no. 6, pp. 940–951, June 2000.

[3] Qing Chen and Thomas R. Fischer, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. on Img. Proc.*, vol. 7, no. 4, pp. 496–505, April 1998.

[4] A. Goldsmith and M. Effros, "Joint design of fixed-rate source codes and multiresolution channel codes," *IEEE Trans. on Comm.*, vol. 46, no. 10, pp. 1301–1312, Oct. 1998.

[5] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

[6] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source-channel coding," *IEEE Trans. Information Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

[7] Min Wang and Thomas R. Fischer, "Trellis-coded quantization designed for noisy channels," *IEEE Trans. Information Theory*, vol. 40, no. 5, pp. 1792–1802, November 1994.

[8] V.A. Vaishampayan and N. Farvardin, "Optimal block cosine transform image coding for noisy channels," *IEEE Transactions on Communications*, vol. 38, no. 3, pp. 327–336, March 1990.

[9] Henrique S. Malvar, *Signal Processing with Lapped Transforms*, Artech House, Boston, MA, USA, 1992.

[10] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The GenLOT: Generalized linear-phase lapped orthogonal transforms," *IEEE Trans. on Sig. Proc.*, vol. 44, pp. 497–507, April 1996.

[11] T. D. Tran, R. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," *IEEE Trans. on Sig. Proc.*, vol. 48, pp. 133–147, Jan. 2000.

[12] Ricardo L. de Queiroz and T. D. Tran, "Lapped transforms for image compression," in *The Handbook on Transforms and Data Compression*, K. R. Rao and P. Yip, Eds., chapter 5. CRC Press, Oct. 2000.

[13] R. Joshi, H. Jafarkhani, J. Kasner, T. Fischer, N. Farvardin, M. Marcellin, and R. Bamberger, "Comparison of different methods of classification in subband coding of images," *IEEE Trans. on Img. Proc.*, vol. 6, no. 11, pp. 1473–1486, 1997.

[14] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," *Proc. IEEE Data Compress. Conf.*, pp. 72–81, March 1997.